JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1997, **67**, 145–159 NUMBER 1 (JANUARY)

CONDITIONED REINFORCEMENT DYNAMICS IN THREE-LINK CHAINED SCHEDULES

BEN A. WILLIAMS

UNIVERSITY OF CALIFORNIA, SAN DIEGO

In two experiments rats were trained on three-link concurrent-chains schedules of reinforcement. In Experiment 1, additional entries to one terminal link were added during one of the middle links to a baseline schedule that was otherwise equal for the two chains, and, depending on the condition, these additional terminal-link presentations ended either in food or in no food. When food occurred, preference was always in favor of the chain with the additional terminal-link presentations (which also entailed a higher rate of reinforcement). When no food occurred at the end of the additional terminal links, the outcome depended on the nature of the stimuli associated with these additional terminal links. When stimuli different from the reinforced baseline terminal links were used for the no-food terminal links, preference was against the choice alternative that led to the extra periods of extinction. When the same stimulus was used for the two kinds of terminal links, preference was near indifference, that is, significantly greater than when different stimuli were used. In Experiment 2, rats learned repeated reversals of a simultaneous discrimination under a three-link concurrent-chains schedule, in which the food or no-food choice outcomes were delayed until the end of the chain. Different conditions were defined by the point in the chain at which differential stimuli occurred. When the middle and terminal links provided no differential stimuli, discrimination was acquired more slowly than when differential stimuli occurred in both links. When differential stimuli occurred in the middle but not the terminal links, acquisition rates were intermediate. Both experiments together show that the effects of stimuli in a chain schedule are due partly to the time to food correlated with the stimuli and partly to the time to the next conditioned reinforcer in the sequence.

Key words: chained schedules, conditioned reinforcement, lever press, rats

---

The status of conditioned reinforcement as an explanatory construct for behavior in chained schedules of reinforcement has remained an issue of controversy. Consider, for example, a three-link chain in which the stimulus in each link of the chain is correlated with its own schedule, access to each successive link of the chain is contingent on responding in the preceding link, and food delivery is immediately dependent on behavior in the terminal link (e.g., with red, green, and blue stimuli correlated with the different links of the chain). In such a three-link chain, the critical issue has been the status of green, the middle-link stimulus, in controlling responding during red and the status of blue in controlling responding during green. The conditioned reinforcement interpretation of chained schedules is that red-key responding is maintained because green has conditioned reinforcement value, which is derived from its own relationship with subsequent events of higher value. The most common version of such an interpretation relies on the concept of the backward transmission of conditioned value via higher order conditioning (for a textbook treatment, see Mazur, 1994). Thus, blue has value because of its contiguous pairing with food, green has value because it is paired with blue, and responding during red is maintained due to the green conditioned reinforcer being contingent on such behavior.

An alternative to the conditioned reinforcement interpretation of chained schedule performance has been advanced by Staddon (1983), in part because of the general finding that it is difficult to maintain initial-link responding in chains with more than two links. Staddon thus argued that the function of the stimuli in the chained schedule was to provide markers for the remaining time to reinforcement. Thus, the behavior maintained by a particular stimulus is due to the time to reinforcement with which it is correlated, not the value of the succeeding stimulus, which is immediately contingent on responding in its presence. Evidence against this interpretation

This research was supported by a grant from the National Science Foundation.

Reprint requests should be addressed to the author, Department of Psychology, University of California San Diego, La Jolla, California 92093-0109 (E-mail: bawilliams@ucsd.edu).

has been provided by Royalty, Williams, and Fantino (1987). Using a chained variable-interval (VI) 33-s VI 33-s VI 33-s schedule, they added a brief unsignaled delay of reinforcement between the response in one link of the chain and the onset of the succeeding link. The rationale was based on the finding that similar delay-of-reinforcement contingencies produce large decrements in behavior (on the order of 70 to 90%) when they are imposed on primary reinforcement contingencies (e.g., Williams, 1976). If behavior during the early links of the chain were maintained by the conditioned reinforcement properties of the stimulus paired with the transition to the next link of the chain, inserting a delay between responding and stimulus onset should produce a decrement in performance similar to that obtained with delayed primary reinforcement. In order to hold the time to food constant, one of the VI 33-s schedules was changed to a VI 30-s schedule with a 3-s unsignaled delay. When this delay contingency was in effect for the advancement from the initial to the middle link of the schedule, response rate in the initial link was decreased by 75 to 80%, with no effect on responding in the middle and terminal links of the chain. Comparable effects on middle-link responding were seen when the delay occurred with respect to the advancement from the middle to the terminal link. The selective nature of the effects of the delay contingency leaves little doubt that the contingency between responding and the onset of the stimulus of the succeeding link of the chain was crucial for maintaining the behavior.

Although the results of Royalty et al. (1987) provide evidence that the response contingency plays a critical role in chained schedule performance, they tell us nothing unequivocally about whether the effects of the stimulus onsets were due to their having acquired conditioned value by higher order conditioning. Perhaps, for example, the contingent stimulus presentations produced a *marking effect,* causing the animal to learn more readily that responding in the early links of the chain was necessary for reaching the end of the chain (cf. Lieberman, McIntosh, & Thomas, 1979). When stimulus onset then was delayed, the behavior that was marked was no longer restricted to pecking the response key, but would include behavior that competes with key pecking. It is also possible that a given stimulus of a chained schedule acquires conditioned value as a function both of its relation to the succeeding link of the chain and of the time to reinforcement correlated with its onset. Given such dual determination, the experimental task is to define the domains of the two different sources of stimulus value.

An attempt to separate the relative roles of conditioned reinforcement and the time to reinforcement signaled by a stimulus was reported by Williams and Royalty (1990). They compared three-link chained schedules in which the middle and terminal links summed in their average time values to 66 s. For one condition, fixed-interval (FI) 33-s schedules were in effect for both the middle and terminal links. For a second condition, VI 33-s schedules were in effect for both links. The critical condition was an interdependent schedule in which the middle-link schedule was a VI 33-s schedule, while the terminal-link schedule was always the time value that resulted from subtracting the particular VI interval that occurred in the middle link from 66 s. The rationale of the comparison was that the chain with the two FI components was equivalent to the interdependent condition in the time to reinforcement signaled by the onset of the middle link (both always 66 s), whereas the interdependent condition differed from the FI condition in terms of the times from the onset of the middle link to the onset of the terminal link (because the middle-link schedule for the interdependent condition was a VI 33-s schedule and that for the FI condition was FI 33 s). Given previous findings that pigeons strongly prefer VI over FI terminal links in simple concurrent chains (e.g., Killeen, 1968), the conditioned reinforcing value of the middle-link stimulus should have been higher in the interdependent condition than in the FI condition, assuming that such reinforcing value was determined by higher order conditioning. In contrast, if the critical variable was time to food signaled by middle-link onset, the FI and interdependent conditions should be equivalent, and both should have had lower initial-link rates than the condition with independent VI 33-s schedules in the middle and terminal links.

The results were complicated by the effect of the initial-link schedule. With FI schedules in the initial link, response rates were consistently highest with independent VI schedules in the middle and terminal links, intermediate with the interdependent schedule, and lowest with FI schedules. With VI schedules in the initial link, response rates were lowest with the FI schedules and were not systematically different for the two types of VI schedules. Thus, overall, the results suggest that time to food signaled by middle-link onset was at best a weak determinant of behavior, whereas the time between middle-link onset and terminal-link onset was a strong determinant.

EXPERIMENT 1

The present study adopts a somewhat different approach to dissociating the effects of time to food and time to the next conditioned reinforcer. To the extent that the value of the middle-link stimulus is determined by its relation to the terminal-link onset, it should be possible to vary the conditioned reinforcement effectiveness of the middle-link stimulus by manipulating the schedule during the middle link. The procedural difficulty is to dissociate such schedule changes from the correlated changes in the times to food from middle-link onset. The approach taken in Experiment 1 was to add additional presentations of the terminal-link stimulus during one of the two middle-link stimuli of a three-link concurrent-chains procedure, but with these additional presentations ending in no food. Such a procedure increases the average time to food as measured from middle-link onset, because only a fraction of the middle-link presentations would be followed by food at the end of the terminal link. In contrast, the procedure decreases the time between the middle-link and terminal-link onsets. However, the additional presentations of the terminal-link stimulus in the absence of food should be expected to decrease its own conditioned reinforcement value, which should counteract, to some degree, the decreased time to the onset of the terminal link. In order to mitigate this effect of terminal-link extinction, the same stimulus was used for both terminal links of a concurrent-chains schedule, so that any decrease in value of the common terminal-link stimulus would affect equally the value of the two different middle-link stimuli.

Rats were presented with a three-link concurrent-chains schedule in which responding was effective only during the initial links. The middle- and terminal-link durations were response independent (to ensure that the obtained time values corresponded to the scheduled values) and the same for the two alternative chains. Initially, the schedule governing the transition from the middle to the terminal links was a variable-time (VT) 30-s schedule for both alternatives of the concurrent chain. The terminal links were always fixed-time (FT) 15-s schedules, and ended with the food reinforcer. The critical manipulation was the addition, in only one of the middle links, of a VT 12-s schedule of presentations of the terminal-link stimulus, but with these additional terminal-link stimulus presentations ending in no food. Thus, the middle link with the additional terminal-link presentations resulted in access to the terminal-link stimulus in a substantially shorter period of time than did the middle link without the added VT 12-s schedule. Because extinction was in effect during these additional terminal-link presentations, however, the predictive value of the middle-link stimulus for food at the end of that chain was substantially reduced. The issue is which of these aspects of the contingency would control behavior.

METHOD

Subjects

Five male albino Sprague-Dawley rats, approximately 5 months old at the start of the experiment, served as subjects. All had prior experimental experience in both a response-acquisition study using delayed reinforcement (see Williams, 1994b, for details) and a successive discrimination problem (Williams, 1994a), both in experimental chambers different from that used here, and with different stimuli. Food deprivation was maintained by allowing 90-min access to laboratory chow approximately 5 min after the end of the experimental sessions (see Hurwitz & Davis, 1983, for the justification of this procedure). Water was continuously available in the home cages at all times. Median body weight at the start of the experiment was 260 g, which increased to 365 g by the end of the study. The subjects were housed in individual cages with a 14:10 hr light/dark cycle.

Apparatus

A custom-built conditioning chamber (24 cm wide by 20 cm high by 26 cm long) was contained within a larger sound-insulating shell, which was equipped with an electric fan for ventilation. The interior chamber was constructed of Plexiglas except for a sheet-metal rear wall and wire-grid floor. The front panel of the chamber was painted black; the remaining walls and ceiling were clear Plexiglas. Mounted on the front panel 11.5 cm above the floor were two nonretractable stainless steel levers, 3 cm in width and protruding 1.8 cm into the chamber. Each lever required a force of at least 0.2 N for operation, with the only feedback for a response being the action of a microswitch connected to the other end of the lever. Directly below each lever and 2 cm above the grid floor was a pellet chute connected to an electromechanical 28-V pellet dispenser (Gerbrands Model G5100) that provided standard 45-mg Noyes chow pellets (improved Formula A). Throughout this experiment, pellets were delivered only to the pellet chute under the left lever.

Mounted 4 cm above the outer edge of each lever was a 28-V miniature light (Sylvania 28 ESB bulb) encased within a recessed bulb holder. Located on the outside of the left side wall was a clicker module (Coulbourn Model E12–05) that presented a 5-Hz auditory clicking stimulus, approximately 80 dB in intensity. Mounted on the rear portion of the ceiling was a 4-ohm speaker through which 80 to 85 dB white noise could be presented. Mounted on the outside of the right wall of the interior chamber was an unshielded 28-V lamp (Bulb 1820) that could illuminate the entire experimental chamber with white light. Experimental events were controlled by electromechanical equipment located in an adjacent room.

Procedure

Because subjects were experienced, they were begun immediately on the baseline procedure involving a three-link concurrent chain, which is shown on the left side of Figure 1. During the initial link of the chain, the two panel lights above the levers were illuminated and a single VI 15-s schedule operated. Reinforcers from the schedule were assigned randomly (p = .5) to the two response levers, and the schedule ceased to time until an assigned reinforcer was obtained. No changeover delay (COD) was used. Instead, response units were defined by a fixed-ratio (FR) value, such that only responses that fulfilled a short FR requirement without interruption (i.e., without the rat switching to the other lever during the sequence) were eligible for reinforcement. Initially the FR value was 1, and it was then increased gradually to 4.

When the schedule requirement during the initial link was fulfilled, responses produced the offset of the front panel lights and the onset of one of the stimuli paired with the middle link, which continued until terminated by a VT 30-s schedule. When the middle link was produced by responses to the left lever, the stimulus present was the houselight illumination located on the outside of the right wall of the chamber. When the middle link was produced by responses to the right lever, the stimulus was the 5-Hz clicking stimulus that emanated from the outside left

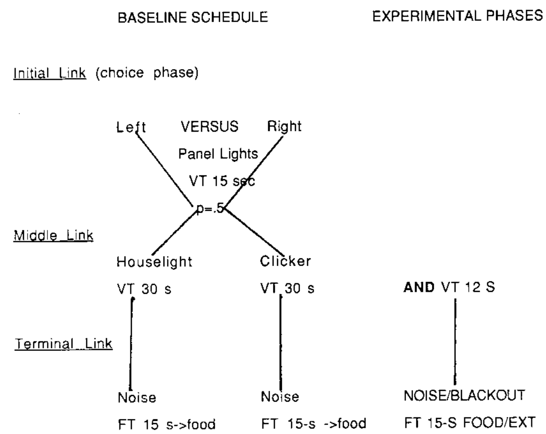

Fig. 1. Outline of procedure for Experiment 1. The procedure shown on the left side operated during the baseline and remained in effect throughout training. The experimental manipulations, shown on the right side, entailed the addition of a VT 12-s schedule during one or the other middle-link stimulus or during both. When the VT 12-s schedule ran concurrently with the VT 30-s schedule, entry to the terminal link was determined by which schedule timed out first. Terminal links produced by the VT 12-s schedule were correlated with either the white noise or a blackout, and could end in food or no food, depending on the experimental condition.

wall of the chamber. Responses during the middle link had no scheduled consequences. The VI schedules in both the initial and middle links were composed of 18 intervals drawn from the exponential distribution specified by Fleshler and Hoffman (1962).

Upon timing out of the middle link of the schedule, the terminal link began with the onset of the 80 to 85 dB noise from the overhead speaker, regardless of whether the left or right lever had been selected in the initial-link choice phase. The terminal link continued until a food pellet was delivered according to an FT 15-s schedule. The initial-link conditions were then reinstated for the next cycle.

After the first 10 sessions in which the schedule values were gradually changed to those used for the remainder of the experiment, the critical procedural manipulation was included (the right side of Figure 1). Now, during one of the two middle-link stimuli (randomized across subjects), an independent VT 12-s schedule was added concurrently with the VT 30-s schedule that otherwise advanced the subject from the middle to the terminal link. Thus, the subject could advance to the terminal link due to either the VT 12-s schedule or the baseline VT 30-s schedule, depending on which timed out first, which meant that the average time to advancement to the terminal link was substantially reduced. In the alternative middle-link stimulus, advances to the terminal link occurred only according to the VT 30-s schedule. In other words, the baseline procedure shown on the left side of Figure 1 continued to operate in addition to the VT 12-s schedule.

The experimental variables were the particular stimuli correlated with the terminal-link presentations produced by the VT 12-s schedule (either the noise or a blackout), and whether or not those terminal links ended in food. The added VT 12-s schedule was presented first with one of the middle-link stimuli, and, when behavior had stabilized, was presented concurrently with the other middle-link stimulus. For the first four conditions, food was presented at the end of all of the terminal-link presentations produced by the VT 12-s schedules. Two of the subjects were initially trained with the blackout as the stimulus for these added terminal-link presentations; the remaining 3 subjects were first trained with the noise stimulus. After stability was achieved, the middle-link stimulus correlated with the VT 12-s schedule was then reversed.

The second four conditions were identical except that food was never presented during the terminal-link presentations produced by the VT 12-s schedule. Food continued to be presented at the termination of the terminal links produced by the VT 30-s schedules, which were always correlated with the noise. The order of conditions (condition designations are shown in Table 1) was 1-2-4-3-7-5-6-8 for Subjects 1, 2, and 5 and 2-1-3-4-8-6-5-7 for Subjects 3 and 4.

In the final condition, the across-condition comparison described above was included within a condition. Now, the VT 12-s schedule operated during both middle-link stimuli and produced a 15-s terminal link that never ended in food (the baseline VT 30-s schedule continued to produce terminal links with food). The difference between the two middle links was that the no-food terminal links produced during one of them was correlated with the same noise stimulus as was used for the regular terminal links, and a blackout was correlated with the other. After stability was achieved, the middle link that produced the noise stimulus during the additional FT 15-s terminal-link presentations was switched, and training continued until stability was again achieved.

Throughout training, experimental sessions were conducted 6 days per week during the lights-on part of the day-night cycle, in the morning and early afternoon hours. A given session terminated after 30 food pellets had been obtained. Training on a given condition continued until a stability criterion was achieved, which required that the medians of the initial-link choice proportions of each block of three sessions differed by no more than ± 0.05 and did not exhibit a monotonic trend over the last nine sessions.

RESULTS

The number of training sessions required to reach the stability criterion ranged from 10 to 23, with a median of 14. There was no trend in the time to reach criterion, either as a function of the order of training or as a

Table 1

Relative response rates in the initial link of the three-link concurrent schedule in favor of the choice alternative that led to the VT 12-s schedule during one of the middle links. The click and houselight designations refer to the discriminative stimuli in the middle link in which the VI 12-s schedule operated. The noise and blackout designations refer to the stimulus in the terminal link produced by the VT 12-s schedule. Data in parentheses are the absolute response rates (responses per minute) summed over both choice alternatives.

Condi-tion	VT 12-s stimulus	Subject 1	2	3	4	5	*M*
Terminal links produced by VT 12-s schedule ended in food							
White noise in terminal link							
1	Click	.55	.61	.62	.65	.45	.58
		(23)	(32)	(23)	(68)	(26)	(34)
2	Houselight	.60	.46	.44	.48	.58	.51
		(37)	(37)	(15)	(67)	(18)	(30)
Blackout in terminal link							
3	Click	.59	.59	.59	.52	.61	.58
		(44)	(43)	(17)	(57)	(34)	(39)
4	Houselight	.52	.57	.64	.67	.63	.61
		(37)	(37)	(15)	(67)	(18)	(35)
Terminal links produced by VT 12-s schedule did not end in food							
White noise in terminal link							
5	Click	.35	.46	.52	.47	.40	.44
		(33)	(35)	(13)	(46)	(14)	(28)
6	Houselight	.56	.48	.42	.45	.64	.51
		(42)	(48)	(11)	(48)	(14)	(33)
Blackout in terminal link							
7	Click	.34	.42	.43	.33	.37	.38
		(43)	(31)	(19)	(32)	(16)	(28)
8	Houselight	.41	.40	.34	.26	.43	.37
		(35)	(33)	(9)	(53)	(14)	(33)

function of the different conditions shown in Table 1.

The first set of conditions, in which food occurred at the end of all terminal-link presentations including those produced by the VT 12-s schedule that operated only during one of the middle links, is shown in the top half of Table 1. The conditions differed only with respect to the stimulus present during the additional terminal-link presentations. Each condition was presented twice, first with the VT 12-s schedule operating during one of the middle links, and then with it operating during the alternative middle link (either during the clicker or the houselight). For all conditions, the choice proportion was calculated as the mean of the relative response rates to the two levers for the nine sessions during which the stability criterion was met.

The preference levels from each condition, in favor of the chain with the added VT 12-s schedule, are shown in Table 1. Averaged over the reversals of the middle-link stimulus correlated with the VT 12-s schedule, preference was .55 when the noise was the extra terminal-link stimulus and .60 when blackout was the terminal-link stimulus. The within-condition standard errors, calculated from the last nine sessions of training, were below .01 for all of the entries shown, indicating very little session-to-session variability. The difference between the stimulus conditions was not significant, $t(4) = 2.24$, $p > .05$. When the choice proportions were combined over Conditions 1 through 4, the average value (.57) was significantly different from the indifference level of .50, $t(4) = 18.23$, $p < .05$.

It should be noted that the VT 12-s schedule had no effect on the relative frequency of food contingent on the two choice responses. This was true because each middle link was entered an equal number of times and each middle-link entry eventually led to food.

However, the VT 12-s schedule did affect the average time between the middle-link onset and the terminal-link onset, and the time between middle-link onset and food at the end of the terminal link. For the middle link with the VT 12-s schedule, the average time to terminal-link onset was 10.6 s, whereas the average time for the middle link without the VT 12-s schedule was 27.2 s. The corresponding times between middle-link onsets and food delivery were 25.6 s and 42.2 s.

The critical data for assessing the role of conditioned reinforcement in the present procedure are the bottom four conditions shown in Table 1. During these conditions, the added VT 12-s schedule occurred during one middle link and then the other, and the FT 15-s terminal-link presentations produced by this schedule never ended in food. For two of these conditions, the stimulus correlated with the terminal links produced by the added VT 12-s schedule was the noise, and for the remaining two conditions the stimulus was a blackout. Note that the noise continued to be correlated with the terminal links produced by the baseline schedules that ended in food.

It is important to recognize that the addition of the VT 12-s schedule substantially decreased the probability of reinforcement for the chain that included the VT 12-s schedule. For this choice alternative, on the majority of occasions, advancement to the terminal link from that middle link resulted in no food, whereas during the alternative chain all terminal-link entries resulted in food. Given that the number of entries into the middle links themselves were equal (because of the random assignment of a single VI schedule to each alternative with a probability of .5), this meant that fewer total food presentations occurred for the chain that included the VT 12-s schedule. The obtained relative frequency of food for the two chain alternatives was .25, averaged over subjects.

The only difference between the bottom two pairs of conditions shown in Table 1 was the presence of the noise or the blackout as the cue correlated with the extinction presentations of the terminal links. All other potentially important aspects of the procedure, including the rate of food presentations for each chain and the temporal distance between the onset of the middle-link stimuli and the terminal link, were similar for the conditions with the different stimuli in the terminal link of the chain. Thus, a comparison between the conditions provides an index of the role played by the noise presentations. The mean difference in preference between the noise and blackout conditions was .10, which was statistically significant, $t(4) = 5.25$, $p < .05$. Thus greater preference occurred when the VT 12-s terminal-link presentations ending in no food were correlated with the noise presentations than with blackout. It also should be noted, however, that both stimulus conditions produced mean preferences for the chain with the added VT 12-s schedule below .50. When blackout was the stimulus correlated with those extra presentations, preference was significantly below .50, $t(4) = 6.31$. When noise was the stimulus, the difference from .50 was not significant, $t(4) = 2.14$.

The obtained pattern of results during the bottom two conditions in Table 1 suggests that there were two separate effects that were superimposed. The significant reduction of preference below .50 seen with the blackout stimulus indicates that the differences in obtained reinforcement rates correlated with the two chains caused a reduction in preference for the chain that had the added terminal-link presentations that ended in no food. Note that this effect is similar, but in the opposite direction, to that shown in the top portion of Table 1, where the additional VT 12-s schedule produced additional food. The second effect is shown by the difference between the corresponding conditions with blackout or noise as the stimulus. Because all other aspects of the procedure were equated, the consistently higher preference with noise as the stimulus shows that noise had some positive value, and that the degree of this value was similar, but in the opposite direction, to the effects of reducing the rate of reinforcement correlated with that chain of the schedule.

The final condition of Experiment 1 combined the bottom two conditions shown in Table 1 into a single condition. Now the VT 12-s schedule operated during both middle links, and all FT 15-s terminal-link presentations resulting from that schedule never ended in food. The only difference was that for one chain, blackout was correlated with its VT 12-s terminal-link presentations, and

Table 2

Relative response rates, calculated from the last nine sessions of each condition, during the initial link when the VT 12-s schedule ran during both middle links, while the no-food terminal-link presentations differed in having the noise or the blackout as the stimulus correlated with extinction. The click and houselight designations refer to the discriminative stimulus in the middle link in which the VT 12-s schedule produced noise as the discriminative stimulus in the terminal link. Preference was calculated in terms of the preference for the noise as the terminal-link stimulus produced by the VT 12-s schedule. In parentheses are the absolute response rates (responses per minute) summed over both initial-link choice alternatives.

Middle-link stimulus	Subject 1	2	3	4	5	*M*
Click	.50	.57	.58	.53	.51	.54
	(42)	(34)	(18)	(35)	(13)	(29)
Houselight	.53	.49	.54	.55	.65	.55
	(41)	(37)	(22)	(36)	(13)	(30)

noise was the stimulus for the other. As can be seen in Table 2, preference was reliably in favor of the chain with noise as the stimulus correlated with the added extinction presentations, $t(4) = 3.95$, $p < .05$.

Tables 1 and 2 also show in parentheses the absolute response rates summed over both levers during the initial links of the schedule. Table 1 reveals that the average response rate for the conditions with all terminal-link presentations ending in food was 34.4 responses per minute, whereas the corresponding rate for the condition in which the terminal-link presentations ended in no food was 29.4. This difference was not statistically reliable, $t(4) = 1.88$. Table 2 shows that similar response rates also occurred during the final condition in which both middle links had the additional terminal-link presentations ending in extinction, but they were not systematically different as a function of the correlated stimulus. It should be noted that the insensitivity of total response rate in the initial link of concurrent-chains schedules to reinforcement variables in the subsequent links has been reported before (Herrnstein, 1964).

Discussion

The results show that the addition of nonreinforced terminal-link entries in the middle link of a three-link chain increased the value of the middle-link stimulus, but also that this increase was counteracted by the decrease in food frequency associated with the nonreinforced terminal links. When blackout signaled the nonreinforced terminal links, a significant preference away from that choice alternative occurred, indicating that the extinction contingency was aversive in character. However, when the noise signaled the nonreinforced terminal links (and also signaled the reinforced terminal links), the level of preference for the additional extinction presentations was near indifference. This reduction in aversiveness of the additional extinction presentations presumably was due to the added value provided by the conditioned reinforcing properties of the noise stimulus.

The failure of the additional presentations of the noise ending in no food to increase preference for that choice alternative above indifference differs from the outcome reported by Williams and Dunn (1991) using pigeons in a two-link concurrent-chains schedule. There the addition of nonreinforced terminal-link presentations contingent on only one choice response increased preference for that alternative, despite the extinction contingency reducing the rate of reinforcement associated with that alternative. Whether this different outcome with two-link chains reflects differences between two- and three-link chains or some other procedural variable (e.g., pigeons vs. rats as subjects) is an unresolved issue.

The present results also stand in opposition to the usual effect of multiple versus mixed schedules in two-link concurrent chains. Such experiments typically arrange two unequal schedules in each terminal link, which are alternately available from trial to trial, with the times to food for the two schedules identical for both choice alternatives. For the mixed alternative the same stimulus is present during the two different schedules; for the multiple alternative different stimuli are correlated with the different schedules. The typical finding has been strong preference for the multiple schedule (Green, 1980). In the present three-link chain, in contrast, the middle link chosen most often was that followed by nondifferential stimuli that were correlated with the two different terminal-link schedules (e.g., the comparison shown in Table 2). This preference for the mixed-schedule alternative was not due to the use of three-link schedules, because a similar preference for the mixed schedule is seen in the results of Williams and Dunn (1991).

One caveat that must be noted regarding the present results is that the size of the effects was notably smaller than that obtained in most studies of preference in concurrent-chains schedules. This does not seem to be due to the insensitivity of preference to different rates of conditioned reinforcement, because a similar low level of preference occurred when the additional terminal-link presentations always ended in food (the first half of Table 1). These conditions produced a relative time to food delivery of .74, whereas the preference for the side with the additional food was only .56. A similar difference occurred between the noise and blackout conditions when the terminal-link presentations produced by the VT 12-s schedule ended in no food. Thus, the low preference levels that were obtained may reflect an inherent insensitivity in a three-link chain in which the terminal-link stimulus is always the same stimulus. Alternatively, the low preference levels may reflect idiosyncratic features of the present procedure (e.g., the use of response units instead of a COD to prevent alternation).

EXPERIMENT 2

The critical finding of Experiment 1 is that preference in a three-link chain was not solely due to the primary reinforcement contingencies. Although these contingencies did control some portion of the behavior, the stimulus schedule was also shown to be an important variable. Thus, the value of the middle-link stimulus was partly determined by the time to food with which it was correlated, and partly determined by its time to the terminal-link stimulus.

Although the results of Experiment 1 were entirely consistent across subjects and across within-condition and between-conditions comparisons, the fact that most of the choice proportions fell in the range of .35 to .65 encourages caution in generalizing the results to other types of chained schedules. It is thus important to specify how the procedure used in Experiment 1 differed from those in other studies. One notable difference was that the primary reinforcement conditions were equal for both choice alternatives, which implies that any differences (e.g., the data seen in Table 2) must be due to differences in the efficacy of conditioned reinforcement. In contrast, most other concurrent-chains studies have involved differences in both primary and conditioned reinforcement. A second notable feature of the the procedure of Experiment 1 was that the different chains chosen by the subjects had the same terminal-link stimulus, in contrast to the differential terminal-link stimuli that normally are arranged.

In order to increase the power of the procedure to distinguish between the conditioned reinforcement and time-to-reinforcement interpretations of the role of chain stimuli, the procedure was changed to a simultaneous discrimination procedure, with only one choice alternative leading to food, but with different links intervening between the two choice responses and their respective delayed outcomes. Presumably, delayed outcomes of choice should retard the rate at which a discrimination is learned. The question posed is how the stimuli that intervene between the choice and outcome modulate this control by the delayed discrimination contingencies. Several different, and sometimes opposing, effects (e.g., conditioned reinforcement, blocking, marking, bridging) have been postulated in different situations (see Williams, 1994b, for a discussion).

Experiment 2 investigated the role of different stimulus sequences in the learning of a simultaneous discrimination modeled after that used in Experiment 1. The major difference was that here one choice alternative led to food and the other did not, and the measure of control was not preference per se but the rate at which the discrimination was acquired. A repeated-acquisition measure involving reversals of the value of the two choice alternatives was used (cf. Williams & Dunn, 1994), such that whenever a learning criterion was achieved, the values of the two choice alternatives were reversed for the next acquisition problem.

The major variable of interest in Experiment 2 was the point in the two chains at which differential stimuli occurred that signaled the differential outcomes at the end of the stimulus sequence. During the nondifferential condition, there was no differential stimulus feedback, because the stimulus sequence following each choice response was always the same (in this case the clicker in the middle link and the noise in the terminal link). Thus, the only way the subjects could learn the discrimination was to utilize the food/no-food trial outcome at the end of the nondifferential stimulus sequence. In the middle-differential condition, the transition from the middle to the terminal links of the S+ versus S− chain (S+ and S− refer to the choice alternatives that led to food and no food, respectively) continued not to have differential stimuli, because the terminal link of both chains was the noise. Now, however, there was differential stimulus feedback upon transition from the initial choice phase to the middle link: Choice of the left lever was always followed by the houselight, and choice of the right lever was always followed by the clicker. Thus, the onset of the differential middle-link stimuli signaled a differential trial outcome but produced the same terminal-link stimulus. Finally, in the differential condition, differential stimulus feedback occurred at each level of chain transition. Choice responses were followed by differential stimuli (the same clicker and light used in Experiment 1 and in the second condition just described). Following choice of the S+ response alternative, the middle link terminated in the noise, and, after an FT of 15 s had elapsed, food was presented at the end of the terminal link. Following choice of the S− alternative, the middle-link stimulus was followed by a blackout, which also terminated on an FT 15-s schedule but ended in no food.

To the extent that the time to food determines the signal value of the intervening stimuli, the prediction is that the differential and middle-differential conditions should produce similar rates of acquisition, because, for both, S+ versus S− choice responses were followed by differential stimuli at the onset of the middle link, which signaled different times to food. The nondifferential condition should produce the slowest rate of acquisition, because no differential stimuli were contingent on the different choice responses. It is nevertheless possible that acquisition still would occur because of control by the delayed primary reinforcement contingencies. Thus, the difference between the nondifferential and middle-differential conditions should provide an index of how a stimulus that signals a particular time to reinforcement facilitates learning over and above whatever control occurs by the delayed food contingency without a signaling effect.

To the extent that the role of the middle- and terminal-link stimuli was determined by higher order conditioning of conditioned value, the prediction is that only the differential condition should produce rapid acquisition, because only during that condition do the differential middle-link stimuli receive differential conditioned reinforcement values because of their pairing with the different terminal-link stimuli, which themselves are correlated with different trial outcomes. The middle-differential condition should be similar to the nondifferential condition, because the concept of backward chaining entails that the value of the middle-link stimuli is the result of its relation to the onset of the terminal link (see Williams, Ploog, & Bell, 1995, for support of this interpretation). Because the terminal-link stimuli were the same in the middle-differential condition, both middle-link stimuli were thus paired with the same conditioned reinforcement outcome, which implies that they should have equal conditioned value. Thus, even though the middle-link stimuli themselves were different for the S+ versus S− choice, their similar conditioned values should strengthen the S+ versus S− choices nondifferentially.

METHOD

Subjects and Apparatus

Six male albino rats, similar to those used in Experiment 1 in terms of experimental histories, weight, and age, served as subjects. Subjects were maintained as described in Experiment 1. The apparatus was the same as Experiment 1.

Procedure

The procedure again was a three-link concurrent-chains schedule of reinforcement, but now with one chain always leading to food and the other ending in no food. During the initial choice link of the chain, the lights above the two levers were illuminated, and transition to the middle link of the schedule was determined by a VI 15-s schedule. Responses had no effect until this schedule had timed out. Then, depending on whether the next response was to the S+ or S− choice alternative, the remainder of the chain was scheduled accordingly, with food at the end of the S+ chain and no food at the end of the S− chain. For both chains, fulfillment of the initial-link requirement resulted in the front panel lights being extinguished and replaced by the appropriate middle-link stimulus.

During the middle links of both chains, a VT 15-s schedule led to an FT 15-s schedule in the terminal links. When the middle and terminal links had been produced by choice of the S+, the terminal link ended in food (a 45-mg Noyes pellet); when produced by choice of the S−, it did not. In either case, the lights above the levers were illuminated and the choice phase of the chain was reinstituted.

Training with one assignment of levers to S+ or S− (e.g., right lever, S+; left lever, S−) continued until 80% of the entries to the middle links were for the S+ alternative, or until 15 sessions of training had occurred. The outcomes correlated with the two levers were then reversed, and training was resumed on the new reversal until the same criterion was reached.

The main independent variable was the stimulus present during each of the middle and terminal links. Table 3 shows those assignments for the 6 subjects. For 3 subjects (R-1, R-2, and R-6), training was first with the middle-differential condition, in which the white noise occurred during both terminal links but with differential stimuli in the middle links. This was followed by training on the differential condition, in which differential stimuli occurred in both the middle and terminal links. Note that for these conditions choice of the left lever always produced the houselight in the middle link of the chain, whereas choice of the right lever always produced the clicker. The third condition was the nondifferential condition, in which the same stimulus sequence occurred in both the middle and terminal links regardless of the choice response (always the clicker in the middle link and the white noise in the terminal link). After two reversals with each of these three conditions, the original pair of conditions was replicated, in reverse order of their initial presentation (replication did not occur for some subjects because of a medical emergency).

Table 3

Stimulus conditions in the middle and terminal links of Experiment 2. Which lever (left vs. right) eventually led to food alternated across problems. The left lever always produced the houselight in the middle link; the right lever always produced the click in the middle link. Three rats were assigned to each of two orders of presentation. However, not all subjects completed all 10 conditions. R-2 received only the first seven conditions; R-4 received only the first nine conditions; R-6 received only the first eight conditions. S+ refers to the chain that led to food in the terminal link; S− refers to the chain that had extinction as the terminal-link schedule.

	Stimulus conditions			
Order	S+ middle	S− middle	S+ terminal	S− terminal
Rats 1, 2, 6				
1	click	light	noise	noise
2	light	click	noise	noise
3	click	light	noise	blackout
4	light	click	noise	blackout
5	click	click	noise	noise
6	click	click	noise	noise
7	click	light	noise	blackout
8	light	click	noise	blackout
9	click	light	noise	noise
10	light	click	noise	noise
Rats 3, 4, 5				
1	click	light	noise	blackout
2	light	click	noise	blackout
3	click	light	noise	noise
4	light	click	noise	noise
5	click	click	noise	noise
6	click	click	noise	noise
7	click	light	noise	noise
8	light	click	noise	noise
9	click	light	noise	blackout
10	light	click	noise	blackout

For the remaining 3 subjects, training initially was with differential stimuli in both the middle and terminal links, followed by training on the middle-differential condition and then the nondifferential condition, followed by a replication of the initial two conditions in reverse order.

Training within a session continued until 50 trials had been presented, or until 1 hr had elapsed. Training within a condition continued for 15 sessions or until at least 80% of the choices on a given session were to the S+ alternative. As shown in Table 3, subjects always received two consecutive reversals on the same condition before a new condition was instituted.

Results

Because different subjects received different numbers of training sessions on the different conditions, some depiction of the data is needed that allows the rate of learning within each condition to be compared. Figure 2 shows one such rendition, by plotting the session number on which a given performance level was achieved against different performance levels, up to 80%, at which point training was terminated for that condition. Note that high values of the sessions-to-criterion measure correspond to slow learning, and low values correspond to fast learning. The results of all reversals (either two, three, or four) with a given condition were averaged together for the results that are presented. Note that the order of presentation for the different conditions was counterbalanced across subjects, with the nondifferential condition in the middle of training, surrounded by different replications of the remaining two conditions.

The results were generally consistent across all subjects, and can be exemplified by Subject R-1. The fastest rate of learning occurred for the condition in which differential feedback occurred in both the middle and terminal links of the schedule (differential condition). Performance when both terminal links were the same noise stimulus (middle differential) was somewhat worse, but was still notably better than when the stimuli in both the middle and terminal links were always the same (nondifferential). Deviations from this pattern occurred primarily in terms of the level of performance in the condition with different middle links but common terminal links (middle differential). For some subjects the pattern was more similar to the condition with nondifferential stimuli throughout the chain (R-1, R-6); for others it was more similar to the condition in which differential stimuli occurred throughout the chain (R-2, R-5, R-4). For all subjects, the ranking of the different conditions was the same when only the last data point on the functions is considered, which corresponds to the number of sessions required to reach the 80% learning criterion.

To provide a better assessment of how the different stimulus conditions permitted learning of the discrimination contingencies, it is helpful to consider the incidence of failures to learn to the 80% criterion within the allotted 15 sessions of training. For the condition with differential stimuli in both the middle and terminal links, there were 22 total reversals presented over subjects, and all 22 were learned to criterion within the allotted sessions. For the condition with differential middle links but common terminal links, there were 20 total reversals, three of which were not learned in the allotted time. One failure occurred for Subject R-6; two failures occurred for Subject R-2. For the condition in which nondifferential stimuli occurred in both the middle and terminal links, a total of 12 reversals were presented, and eight of these failed to be learned to criterion within 15 sessions. At least one failure occurred for all 6 subjects (usually the first reversal of the pair), and two failures to meet the learning criterion occurred for 2 subjects (R-2 and R-6).

Discussion

The role of stimuli that intervene between the choice response and its delayed outcome is best seen by the comparison between the nondifferential condition, in which the same stimuli were present in the middle and terminal links of both chains, and the differential condition, in which different stimuli were present in both the middle and and terminal links. In the latter case, rapid learning occurred, in that all subjects reached the 80% criterion in only a few sessions. For the nondifferential condition, in contrast, 8 of the 12 exposures to the condition produced failures to meet the learning criterion, and for the remainder learning was substantially retarded. Thus, having differential stimuli seems to be critical to achieving good control by delayed outcome contingencies in a simultaneous discrimination.

The critical theoretical question is how the differential stimuli in the middle and terminal links of the chain achieve the effect of facilitating discrimination. Traditional interpretations of chained schedule performance have invoked the idea of backward transmission of value as the underlying mechanism. That is, the occurrence of food at the end of one chain, but not after the other, transmits differential value to the different stimuli in the preceding terminal links, which in turn transmit differential value to the different

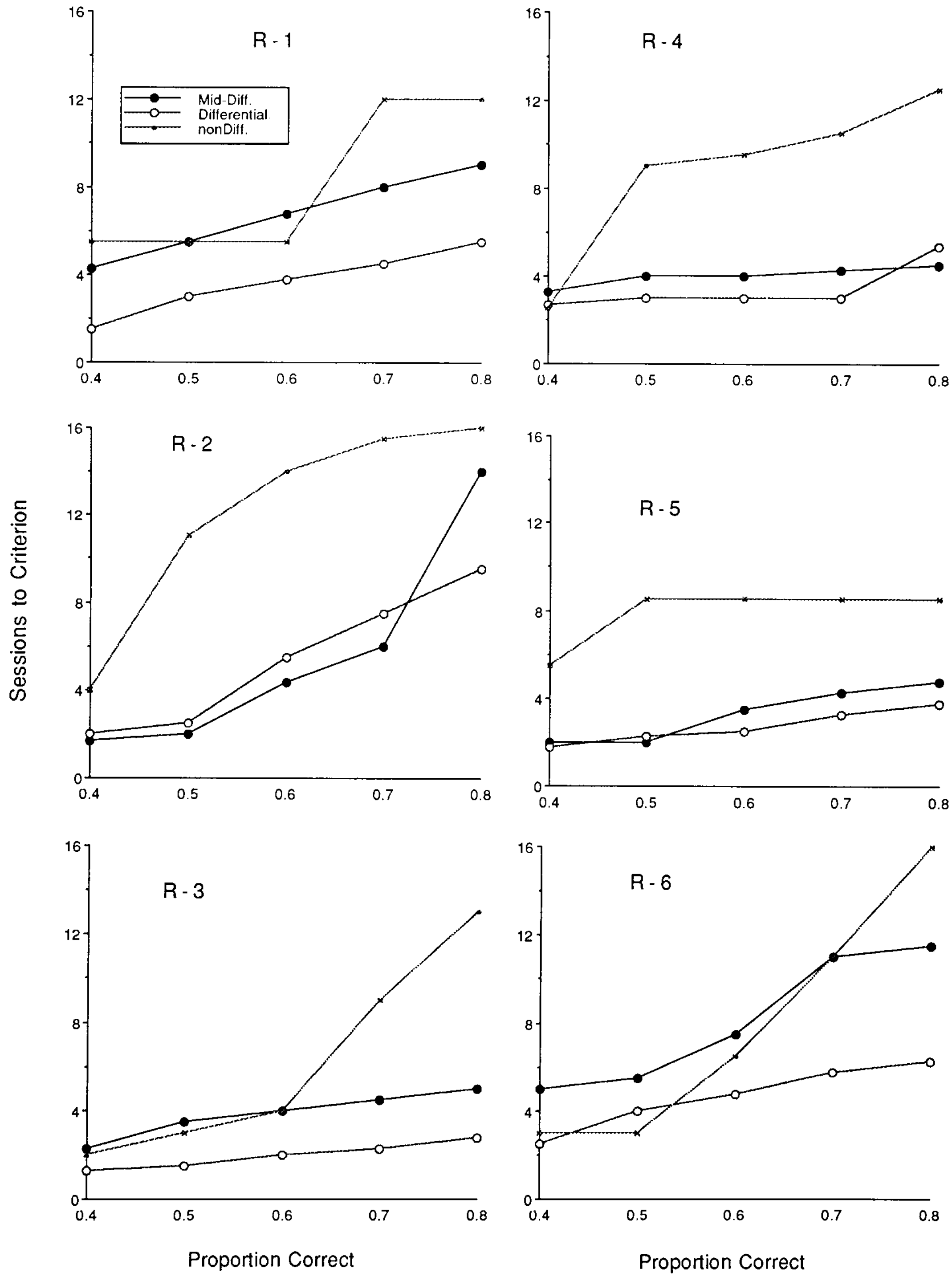

Fig. 2. Acquisition data for individual rats in Experiment 2. Plotted on the ordinate are the number of training sessions required to reach various levels of discrimination proficiency; hence, smaller numbers correspond to more rapid learning. The different conditions for each subject correspond to the differential, middle-differential, and nondifferential stimulus conditions. The specific stimulus sequence presented to each subject is shown in Table 3.

stimuli in the middle links. When the two chains do not involve differential stimuli, however, there is no way the differences in value at the end of the chain can be transmitted differentially to earlier links of the chain. Accordingly, therefore, the absence of any differential stimuli should (and did) produce very poor learning.

The most interesting condition from a theoretical perspective is that in which differential stimuli occurred in the middle links of the chains but not in the terminal links. In general, performance on this condition was substantially better than when the middle-link stimuli were also nondifferential, but was also often notably worse than when differential stimuli occurred in both the middle and terminal links. The former finding is most problematic for the idea of backward transmission of value, because the common white noise stimulus in both terminal links should have possessed an average value for that stimulus, which was then transmitted nondifferentially to the preceding clicker and houselight middle-link stimuli. Thus, performance should have been similar to that when both the middle- and terminal-link stimuli were nondifferential. The superiority of the middle-differential condition over the nondifferential condition thus implies that something other than, or in addition to, the backward transmission of value is involved.

One possible source of the beneficial effects of the differential middle-link stimuli in the middle-differential condition is that their onsets were discriminative cues for their respective times to food. As argued by a variety of investigators (e.g., Staddon, 1983; see Williams, 1994b, for a review), the value of a stimulus is determined by the time to reinforcement that it signals, with no role played by other stimuli that intervene between the stimulus and reinforcer. The finding that the differential condition was consistently superior to the middle-differential condition suggests that this account is insufficient to explain the present data. That is, the information value of the middle-link stimuli was similar for the differential and middle-differential conditions, but substantially different rates of learning occurred. It should be noted that the time-to-reinforcement view of stimulus function fails to explain a variety of other data as well (Royalty et al., 1987; Williams & Royalty, 1990).

Although time to reinforcement signaled by the middle-link stimuli is inadequate to explain the present set of results, it may play some part in the complete explanation. That is, the conditioned reinforcement properties of the middle-link stimuli could be due in part to the time to food signaled by their onset and by the time to the terminal-link stimulus. Such a dual-process view is consistent with a variety of other data (e.g., Williams, 1994b), including those presented in Experiment 1.

An alternative view of the facilitative effects of the middle-differential condition is that the onset of the nondifferential terminal-link stimulus following the differential middle-link stimuli created the possibility of conditional discrimination. Thus, just because the same terminal-link stimulus nominally occurred on both S+ and S− trials did not necessarily mean that the different presentations of the noise did not have differential value to the subject. That is, white noise after the clicker could have functioned as a different stimulus event than white noise after the houselight. The poorer performance with the middle-differential condition relative to the differential condition might then be explained by the difficulties involved in establishing and maintaining the conditional discrimination. The absence of any response requirements in the middle and terminal links in the present study prevents the assessment of whether such a conditional discrimination was actually formed. However, other studies of conditioned reinforcement (e.g., Williams & Dunn, 1994) have shown clearly that acquisition of conditional discrimination contingencies may be involved. Given that possibility, the backward transmission of value as a complete explanation of chain schedule performance cannot be excluded.

REFERENCES

Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 5,* 529–530.

Green, L. (1980). Preference as a function of the correlation between stimuli and reinforcement outcomes. *Learning and Motivation, 11,* 238–255.

Herrnstein, R. J. (1964). Secondary reinforcement and rate of primary reinforcement. *Journal of the Experimental Analysis of Behavior, 7,* 27–36.

Hurwitz, H. M. B., & Davis, H. (1983). Depriving rats of food: A reappraisal of two techniques. *Journal of the Experimental Analysis of Behavior, 40,* 211–213.

Killeen, P. (1968). On the measurement of reinforcement frequency in the study of preference. *Journal of the Experimental Analysis of Behavior, 11,* 263–269.

Lieberman, D. A., McIntosh, D. C., & Thomas, G. V. (1979). Learning when reward is delayed: A marking hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes, 5,* 224–242.

Mazur, J. A. (1994). *Learning and behavior* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Royalty, P., Williams, B. A., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior, 47,* 41–56.

Staddon, J. E. R. (1983). *Adaptive learning and behavior.* Cambridge: Cambridge University Press.

Williams, B. A. (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 26,* 441–449.

Williams, B. A. (1994a). Blocking despite changes in reinforcer identity. *Animal Learning & Behavior, 22,* 442–457.

Williams, B. A. (1994b). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review, 1,* 457–475.

Williams, B. A., & Dunn, R. (1991). Preference for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 55,* 37–46.

Williams, B. A., & Dunn, R. (1994). Context specificity of conditioned reinforcement effects on discrimination acquisition. *Journal of the Experimental Analysis of Behavior, 62,* 157–167.

Williams, B. A., Ploog, B. O., & Bell, M. C. (1995). Stimulus devaluation and extinction of chain schedule performance. *Animal Learning & Behavior, 23,* 104–114.

Williams, B. A., & Royalty, P. (1990). Conditioned reinforcement versus time to reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior, 53,* 381–393.

Received October 30, 1995
Final acceptance September 30, 1996